

DOCUMENT RESUME

ED 117 199

95

TM 005 059

AUTHOR Sherman, Thomas M.
TITLE Formative Student Evaluation of Instruction. TM Report 46.
INSTITUTION ERIC Clearinghouse on Tests, Measurement, and Evaluation, Princeton, N.J.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO ERIC-TM-46
PUB DATE Oct 75
CONTRACT NIE-C-400-75-0015
NOTE 15p.
AVAILABLE FROM ERIC Clearinghouse on Tests, Measurement, and Evaluation, Educational Testing Service, Princeton, N.J. 08540 (free while supplies last)

EDRS PRICE MF-\$0.76 HC-\$1.58 Plus Postage
DESCRIPTORS Effective Teaching; Elementary Secondary Education; Feedback; *Formative Evaluation
IDENTIFIERS *Student Evaluation of Teacher Performance

ABSTRACT

Current practices and controversial issues pertaining to student evaluation of instruction are briefly reviewed and followed by a discussion of the application of formative evaluation to this endeavor. Two reasons are generally given for conducting teacher evaluations by students: identification of good teachers and instructional improvement. Summative evaluation appears to provide the identification but does not appear to improve instruction. The use of summative evaluation alone ignores the dynamic and process characteristics of teaching and makes teaching seem to be a static product which is unchanging and unidimensional. A formative approach appears best suited to provide information on improving instruction and should meet these three criteria: (1) it should be specific to a lesson, objective, or teaching behavior; (2) it should be conducted frequently; and (3) it should yield information which may foster the improvement of instruction during the teaching process. Such evaluation must be an integral part of the teaching process. In this way a closed feedback loop is created wherein information may be continually recycled. Formative evaluation should be embedded in a series of comprehensive and systematic procedures designed to evaluate instructional quality. Using formative evaluation by students effectively and in concert with other forms and sources of information should help teachers answer many questions about their teaching effectiveness and ultimately result in teaching improvement. (RC)

TM REPORT 46

OCTOBER 1975

FORMATIVE STUDENT EVALUATION OF INSTRUCTION

Thomas M. Shernan

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

Abstract

Current practices and controversial issues pertaining to student evaluation of instruction are briefly reviewed and followed by a discussion of the application of formative evaluation to this endeavor. Two reasons are generally given for conducting teacher evaluations by students: identification of good teachers and instructional improvement. Summative evaluation appears to provide the identification but does not appear to improve instruction. The use of summative evaluation alone ignores the dynamic and process characteristics of teaching and makes teaching seem to be a static product which is unchanging and unidimensional. A formative approach appears best suited to provide information on improving instruction and should meet these three criteria: 1) it should be specific to a lesson, objective, or teaching behavior; 2) it should be conducted frequently; and, 3) it should yield information which may foster the improvement of instruction during the teaching process. Such evaluation must be an integral part of the teaching process. In this way a closed feedback loop is created wherein information may be continually recycled. Formative evaluation should be embedded in a series of comprehensive and systematic procedures designed to evaluate instructional quality. Using formative evaluation by students effectively and in concert with other forms and sources of information should help teachers answer many questions about their teaching effectiveness and ultimately result in teaching improvement.

Introduction

The use of student evaluation of instruction has generated not only a great deal of controversy but also a large volume of research. Although much of the resulting data and journalistic comment is contradictory, there appears to be fairly wide agreement that students should evaluate instruction and that their judgements are likely to be of some value. Much of the work in this area has yielded well-constructed evaluation instruments that have both validity and reliability. Yet, even with well-constructed instruments, there is no universal agreement as to the best procedure for gathering student opinions about instruction. This paper will briefly review some of the controversial issues connected with student evaluation of instruction. First, current practices in student evaluation of instruction will be outlined; then the application of formative evaluation techniques to student evaluation of instruction will be discussed. This paper will not provide a comprehensive review of student evaluation of instruction; rather, it will focus on the use of formative evaluation procedures for student evaluation of instruction and explore some of the implications of this approach.

Determining the Purposes of Evaluation

The major reason for collecting evaluation data of any sort is to provide the necessary information for decision making. Data may be used for making decisions about whether a product is good, bad, or in need of revision. In addition, the information that is gathered may be studied to determine its quality. Since specific decisions are usually made according to the functions of a product or process, the information that is gathered must be oriented to the needs of the decision maker.

Bloom, Hastings, and Madaus (2) state that there are at least two ways of looking at evaluation depending upon the purpose of the evaluation or the nature of the decisions to be made. They define summative evaluation as assessment "at the end of a course or topic or unit; that is, when no subsequent changes in treatment for that learning will be made" (p. 262). Summative evaluation provides a "general assessment of the degree to which the larger outcomes have been attained over the entire course or some substantial part of it" (p. 61). On the other hand, these authors describe formative evaluation as a procedure used "during a course, when (presumably) changes can be made in the transactions of subsequent instruction on the basis of current attainment" (p. 262). Formative evaluation yields information to "determine the degree of mastery of a given learning task and to pinpoint the part of the task not mastered" (p. 61). Distinctions between formative and summative evaluation are also based on the timing of evaluation and the level of generalization. Formative evaluation is undertaken much more frequently and is more specific to definable tasks or objectives than summative evaluation, which is oriented toward broader goals.

Summative evaluation, then, measures broad abilities or conceptualizations, whereas formative evaluation focuses on specific components of these larger abilities. In essence, summative evaluation provides information about the overall product of a process once that product has been developed; formative evaluation delivers information about a product or process while the product is being developed or formed.

Evaluating teaching as a Process

Although Bloom, Hastings, and Madaus and others (35, 36) discuss formative and summative evaluation in terms of educational products and their effect on producing learning, the terms are also applicable to teaching when teaching is viewed as an educational product intended to produce learning. Though teaching is in many senses an instructional product, it differs from instructional products such as programmed materials and textbooks in that it is not so easily evaluated and revised. This difficulty arises because teachers' performances are never exactly the same; that is, a teacher approaches the same lesson differently each time even when the content is the same. Therefore, identifying effective and ineffective components of a specific teaching performance is often difficult or impossible. This problem, combined with the difficulties presented by the broad range of student responses to a teaching performance, makes teaching a very complex process to evaluate.

The teaching process consists of many elements. For example, Gagne and Briggs (16) describe a set of eight steps in lesson preparation and presentation. Each step represents a large set of teaching behaviors necessary for producing a good lesson. In addition, courses are made up of a series of these lessons, and, just as each of Gagne and Briggs' eight steps contribute to the development of a lesson, so each lesson contributes to the success or failure of a course. The degree to which each lesson contributes to the success or failure of student learning can only be determined as each lesson is delivered. This requires evaluation of the teaching process for each lesson so decisions can be made by the instructor about changes in lessons and courses.

In addition to the course as a whole, then, each lesson should be evaluated as a separate unit. Summative evaluation of instruction can yield information for comparing the general performance of one teacher with the performance of others; however, without corresponding information about each lesson, the teacher cannot make decisions about specific changes to improve instruction. Summative evaluation may provide data indicating that one course was more poorly taught than another or that students in one section did not learn as much as students in another section of the same course; however, when this information becomes available, it is usually too late to change methods, content, or procedures. Formative evaluation can provide the necessary information about specific

actions a teacher engaged in at specific points; this kind of information helps the teacher make decisions about adjustment of content, methods, or procedures.

One final note: Summative evaluation used exclusively treats "good teaching" as a trait of the teacher rather than a developable skill. According to Cohen, Trent, and Rose (9), "Such a procedure permits the prediction of teacher effectiveness, not in terms of a set of absolute characteristics but in terms of the traits regarded as significant by judges of differing educational philosophies, differing personal preferences and in differing situations" (p. 1042). This situation is analogous to what has occurred in education for centuries. Intelligence has been treated as a trait, and evaluation has been used to select "good learners" and screen out "poor learners." Summative evaluation alone, may seem to indicate that "good teaching" approximates a normal distribution in a similar manner. Implied operational definitions suggesting that good teaching is whatever the questionnaire measures can initiate much misguided behavior on the part of students, faculty, and administrators. Many of these problems have been pointed out by Dansereau (11), Kerlinger (22), and Marshall (26).

Interpreting the Data: Some Cautions

Recent reviews of student evaluation of instruction (10, 40) present scores of studies, many of which have contradictory results; however, these reviews show that some general trends have emerged.

Costlin, Greenough, and Menges (10) note that student ratings of teacher effectiveness "are positively correlated with the teacher's 'agreeableness,' 'emotional stability,' and 'enthusiasm'" (p. 531). In addition, Trent and Cohen (40) state that "the consensus of relatively plentiful research indicates that student evaluation...provides useful and reliable information about at least three aspects of college teaching: 1) the skill of a teacher in terms of his personal effectiveness; 2) the rapport between the teacher and his students; and, 3) the organization and management of a particular course" (p. 1047).

The vast majority of studies report a similar pattern in the way evaluation of instruction data was gathered. That is, at the end of a course, students were given a questionnaire containing items intended to assess teaching quality in relation to factors such as "Student Accomplishment," "Teacher's Presentation," and "Organization" (27). Students then completed the questionnaire by rating their instructors on these factors using a Likert-type scale. The results were then tabulated and summarized to provide an overall picture of teacher effectiveness. Almost always, the questionnaires were constructed and administered by someone not associated with either the teacher or the course. This procedure fits well into the definition of summative evaluation.

Generally, two reasons are given for conducting these teacher evaluations: The first is to identify good teachers; the second is to improve the quality of instruction (14, 21, 42). Expressions of caution about interpretation of summative evaluation data are commonly found in the literature (18, 22). How well this approach actually identifies good teachers is a controversial question (33), though there is some apparent agreement that summative student evaluation of instruction is at least one legitimate procedure for identifying good teachers (10). One common reason often given for identifying good teachers is to reward those teachers with salary increments and tenure. Some authors appear to be cynical about the effectiveness of using student ratings in this way. For example, Zelby (44) states that student evaluation of faculty "in almost any form will become widely and rapidly accepted because it will permit academic administrators to shirk the responsibility of exercising judgement in the evaluation of teaching performance, and at the same time to use (student evaluation) as tangible proof that something is being done about improving teaching" (p. 1270). Apparently, no evidence is available that good teachers identified through student evaluation actually are rewarded by administrators, though evaluation data from students does appear to be increasingly included in personnel decisions.

Kerlinger (22) lists six deleterious effects often connected with the common student evaluation procedure:

- * It alienates professors from their work.
- * It creates instructor hostility.
- * It erodes professional motivation.
- * It undermines professional autonomy.
- * It erodes professional responsibility for instruction.
- * It undermines student respect for instruction.

In addition, Kerlinger notes that "Department chairmen and deans evaluate instructors, but rarely to improve instruction. Such evaluation is geared, rather, to administrative ends..." (p. 354).

Since summative evaluation is global and non-specific and thus provides little or no data for revising specific procedures, it is not surprising that reports about the effect of student evaluations on instructional improvement have been contradictory. As Centra (7) noted, "...an overall judgement does not give a teacher the kind of specific information needed for improvement" (p. 12).

Thus, a summative approach to student evaluation of teaching may provide some information about who is and who is not a good teacher. This kind of information fulfills the identification

purpose of student evaluation. However, the second function — instructional improvement — does not appear to be well served by a summative approach. The use of summative evaluation alone ignores the dynamics of the teaching process and makes teaching appear to be an unchanging product that can simply be characterized as good or bad. On the other hand, McKeachie, Linn, and Mann (28), have indicated that teaching involves a number of complex interactions and that "When we ask, 'Which teachers are most effective?' we need to add further, 'For which objectives?' and 'For which students?'" (p.444).

Formative Approaches to Student Evaluation of Instruction

Several studies have been conducted in recent years in which feedback was given to instructors; these studies all employed a similar methodology (3, 5, 6, 19, 29, 30, 31, and 34). An evaluation form similar to the forms used for summative evaluation was given to students about halfway through a course. The results of this administration were then passed on to the instructor, and the evaluation was administered again at the end of the course. Differences were then noted between instructors who received feedback and those who did not. Reports on the results have been contradictory. Miller (29) found positive change in only one of three classes; Bultman (5) found feedback did not produce any change; and Oles and Lencoske (30) found change that was the opposite of what was expected. On the other hand, Centra (6), Braunstein, Klein, and Pachla (3), Pambookian (31), and Rous, *et al.* (34) noted change among instructors who received feedback.

One reason for the differing reports may be that the feedback in the studies in which change occurred was qualitatively different. Centra (7) noted that to be effective, feedback from student ratings should be provided in the context of some comparative data. This point was also made by Kerlinger (22) and Yonge and Sassenrath (43). Another important factor in teacher change based on student evaluations appears to be the teacher's self-perception. Both Pambookian (31) and Centra (6) noted change among teachers who rated themselves better than their students rated them. Apparently the discrepancy between self-perception and the perception of students, if not too great, serves as a motivator of change in the desired direction. Finally, Centra (7) and Pambookian (31) suggest that teachers should be provided with information on how their instruction may be improved. There does appear to be significant evidence that feedback from student ratings can promote some change in subsequent ratings if the feedback is delivered to the instructor properly.

Some Criteria for Formative Evaluation

In the studies cited above, evaluation was neither conducted frequently enough nor was it specific enough to qualify as formative evaluation. In relation to the two major purposes for conducting

student evaluation, a formative approach appears best suited to provide information for improving instruction. Centra (7) notes that "when student ratings of instruction are used formatively — that is, when they are used by instructors as a source of feedback on their teaching — the evidence indicates that some changes are made by the instructor" (p. 16). Grasha (17) states that "the real test of the effectiveness of a teacher's behavior is how the things he or she did were seen as useful contributions to the students' experience with the course" (p. 2). Similarly, Brightman (4) states that "for instructional evaluation to be performed most meaningfully, it must be conducted in terms of specific education objectives..." (p. 33).

Formative evaluation of instruction by students, then, should meet the following criteria: (1) It should be specific to a lesson, objective, or teaching behavior; (2) it should be conducted frequently; and, (3) it should yield information that can foster the improvement of instruction during the process of instruction. Such evaluation, by its nature, must be an integral part of the process of teaching. In this way, a closed feedback loop is created wherein information may be continually recycled.

A report by Sherman and Winstead (38) provides the only description of a procedure for formative student evaluation of instruction. The authors asked students to use a Likert-type scale to rate (a) the value of the class for personal gain, and (b) the quality of instruction at the end of each class. The data was used for two purposes: First, it allowed poor classes to be identified, and, second, the ratings allowed identification of individual students who did not like the instructional style of the course or did not find the course valuable.

The data was gathered by having each student record the ratings on an individual form that was kept in a personal file folder. Each student received his or her folder prior to each class and returned it following the class. The authors reported that the rating process required 10 to 30 seconds once the students were accustomed to the routine. High reliability was reported, and the authors stressed the problem-free nature of the procedure. Thus, the procedure, though simple, did provide for systematic collection of formative student evaluation data.

Sherman and Winstead present several examples of the ways the data was used. They state that "Formative evaluation carried out on a session-by-session basis has the potential advantage of behavioral specificity. This specificity can facilitate changes in instructional programs because of the immediacy and specific nature of the feedback received from students" (p. 38).

It appears that the routine use of student formative evaluation of instruction is not widespread. The notion of formative evaluation has been discussed for a number of years, as noted by

Scriven (36), and currently there is much emphasis on improvement of instruction; thus, it is somewhat perplexing that this form of evaluation has not been more widely employed. Perhaps one reason is that the vast majority of such evaluation efforts (most of which are unreported in the literature) are quite haphazard. Also, as a study by Marshall (26) suggests, both faculty and students may view the evaluation process with suspicion. Another problem may be the resistance of some and inability of others to define "good teaching." Certainly, numerous commentators have discussed the difficulty of identifying effective teaching. Perhaps, as Popham (32) concluded, "the quality of learning which transpires in a given instructional situation is a function of particular instructional procedures employed by a particular instructor for particular students with particular goals in mind" (p. 2). Finally, the absence of practical procedures to gather and utilize formative evaluation data has probably inhibited many instructors.

The Need for Research and Training

Several important issues should be confronted in future work on formative student evaluation of instruction. First, practical procedures should be developed and field tested in a wide variety of educational settings. Second, research should be conducted on the effect of formative student evaluation of instruction on instructors, students, and student achievement. Third, teachers should be trained to conduct and use formative evaluation. Fourth, rating procedures should be thoroughly examined. And, finally, student formative evaluation should be placed within a framework of broadly based evaluation of instruction.

Certainly formative student evaluation of instruction will not be widely employed unless systematic procedures for conducting such evaluation are available. At present, only one systematic procedure has been reported. Other approaches could employ feedback procedures such as requesting students to smile and be attentive when certain teaching behaviors are engaged in or absent (23). Computer hardware could also be employed to provide a cumulative record of teacher behavior as perceived by students (7). Hohen (20) and Fox, Luszki, and Schmuck (15) present a series of evaluation forms that could be adapted to gather formative student evaluation data. The development of appealing procedures for gathering information should increase the popularity of this form of evaluation.

There is no direct evidence presently available on the effect of formative feedback from students on instructor behavior. Evidence from summative studies, however, tends to indicate that the form of the data gathered may have an impact on how or whether it is used to alter instruction. For example, Centra (7) found that teachers who received no comparative information to aid in data interpretation did not change their teaching. Also, Sherman (37) has reviewed some preliminary evaluation results that suggest that

It is important for the data to be specific. However, there appear to be no studies that document actual teachers' behavior changes resulting from student feedback data; most studies report only changes indicated by students' re-evaluations.

It would be interesting to note the parameters of any changes that did occur. For example, Marsh, Fliener and Thomas (25) found that offering instructors feedback led to improved subsequent ratings; however, the achievement scores of students whose feedback was offered did not differ from the scores of students whose feedback was withheld. Thus, instructors who received feedback apparently modified their behavior enough to improve their student ratings, but these changes did not produce increased achievement.

One other factor that could interact with formative student evaluation of instruction is the teacher's self-perception. It may be possible that even teachers with wide discrepancies between their self-perception and their students' perceptions would change if they were provided with formative data. Attempts at improvement based on specific formative evaluation data may be far less discouraging than attempts based on data from global measures.

Some teachers may not view teaching as a complex process; that is, they may claim that "teaching is teaching." Others may be content with their current practices no matter what the results. In addition, the continuing introduction of new theories and concepts of instruction (1, 12, 13) may leave many instructors somewhat behind the current cutting-edge of innovation in instruction. Miller (29) and Webster and Mendro (41) among others have noted that an instructor's desire to change may be an important factor in any change that does occur. Since the purpose of formative evaluation is to promote and guide change, those who do not view teaching as a dynamic process may not recognize either the need for change or the value of formative evaluation data. To remedy this situation, instructors may have to be familiarized with newer instructional practices and the purposes of newer evaluation procedures. Since formative evaluation should be designed to provide the decision information each individual instructor requires, instructors who are willing to change should find formative student evaluation of instruction valuable. The task of orienting teachers to these procedures is a major challenge for those in a position to foster the effective use of evaluation data.

The role of the rater and rating procedures may be the area most in need of careful study. For example, it is quite possible that students are such poor judges of the quality of instruction that, statistical reliability notwithstanding, student ratings may be of no actual value. Yonge and Sassenrath (43) and Centra and Linn (8) emphasize the importance of the context in which ratings are interpreted. Tolor (39) also states that his findings suggest "that students are indeed quite inaccurate in their perceptions of poor teachers" (p. 103). McKeachie, Linn, and Mann (28) conclude

that the major slippage in validity studies was due to differing goals of teachers and students. Cohen, Trent, and Rose (9) state that "In order to predict how a particular observer or even sets of observers will judge a given teacher, it is necessary to discover first what the observer's expectations are with respect to the teacher role" (p. 1043). Sherman (37) asked students to give reasons for the numerical ratings they gave a teacher at the end of each class and found that many students were unable or unwilling to do so.

Thus, students may need to be trained to judge and record judgements of good teaching, particularly if the judgements are to provide data for improving instruction. It is also unclear whether the most helpful ratings utilize numbers in a Likert-type scale, some other form of rating scale, or narrative comments made by students. In any event, there is the possibility that the format of the instrument will significantly affect the outcome. Moreover, as Centra (7) pointed out, it seems clear that students may not be very discriminating judges since "ratings are typically biased in a positive direction" (p. 5). Students trained to judge instruction may provide much more useful formative data than untrained judges.

Finally, formative student evaluation of instruction should be undertaken within a framework of comprehensive and systematic procedures designed to evaluate the quality of instruction. That is, formative student evaluation should be one component in a series of evaluation procedures. Supporting this idea, Costin, Greenough, and Menges (10) state, "We wish to emphasize that student ratings of undergraduate teaching fall far short of a complete assessment of an instructor's teaching contribution." In addition, Grasha (17) notes that evaluation must be broadly based. Since teaching is a complex process, the evaluation of teaching must reflect this complexity. Attention must be given to planning, content, objectives, assessment procedures, and so forth, and consideration should be given to both formative and summative evaluation. Sanders and Cunningham (35) and Lawson (24) have presented comprehensive schemata for implementing formative evaluation. According to these models, teachers should employ students' formative evaluations as well as formative evaluations by others for all components of instruction. A total evaluation effort should also include summative evaluation. Sanders and Cunningham (35) state, "It seems clear that formative evaluation is not an undisciplined undertaking and that inquiry methods currently used in a wide range of disciplines need to be applied, when relevant, in answering formative evaluation questions" (p. 233). Thus, the more data gathering methods applied and sources sought, the more decision information is made available to instructors for improving the quality of the instructional process. Using formative evaluation by students effectively and in concert with other forms and sources of information should help teachers answer many questions about the effectiveness of their teaching and ultimately result in an improvement in all phases of their teaching.

REFERENCES

1. Bloom, B. S. Learning for mastery. Evaluation Comment, University of California at Los Angeles, Center for the Study of Evaluation of Instructional Programs: May 1968, 1.
2. Bloom, B. S., Hastings, J. T., & Madaus, G. F. Handbook of formative and summative evaluation of student learning. New York: McGraw-Hill, 1971.
3. Braunstein, D. N., Klein, G. A., & Pachla, M. Feedback expectancy and shifts in student ratings of college faculty. Journal of Applied Psychology, 1973, 58, 254-258.
4. Brightman, R. W. Strategies for change, part IV: comparing instructional effectiveness. Mesa, Calif.: Coast Community College District, 1971. ED 050 712.
5. Bultman, J. E. Concerted effort can change the teacher image. Journal of Experimental Education, 1972, 40, 43-45.
6. Centra, J. A. Effectiveness of student feedback in modifying college instruction. Journal of Educational Psychology, 1973, 65, 395-401.
7. Centra, J. A. The student as godfather? the impact of student ratings on academia. Paper presented at the First Invitational Conference on Faculty Effectiveness as Measured by Students, Temple University, Philadelphia, Pa., 1973. ED 079 338.
8. Centra, J. A. & Linn, R. L. Student points of view in ratings of college instruction. Princeton, N.J.: Educational Testing Service, 1973. ED 089 581.
9. Cohen, A. M., Trent, J. W., & Rose, C. Evaluation of teaching. In R.M.W. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand McNally and Co., 1973. Pp. 1041-1052.
10. Costin, F., Greenough, W. T., & Menges, R. J. Student ratings of college teaching: reliability, validity, and usefulness. Review of Educational Research, 1971, 41, 511-535.
11. Dansereau, R. A. Some myths on student evaluation of faculty. Improving College and University Teaching, 1973, 21, 45-50.
12. Davies, I. K. Competency based learning: technology, management, and design. New York: McGraw-Hill, 1973.
13. Davis, R. H., Alexander, L. T., & Yelon, S. L. Learning system design: An approach to the improvement of instruction. New York: McGraw-Hill, 1974.

14. Dillman, T. When college students grade the faculty. Today's Education, 1970, 59, 62-71.
15. Foy, R., Luszki, M. B., & Schmuck, R. Diagnosing classroom learning environments. Chicago: Science Research Associates, 1966.
16. Gagne, R. M. & Briggs, L. J. Principles of instructional design. New York: Holt, Rinehart and Winston, Inc., 1974.
17. Grasha, A. F. Evaluating teaching: some problems. Cincinnati, Ohio: Institute for Research and Teaching in Higher Education, Cincinnati Univer., 1972. ED 071 582.
18. Harari, O., & Zedeck, S. Development of behaviorally anchored scales for the evaluation of faculty teaching. Journal of applied Psychology, 1973, 58, 261-265.
19. Heyn, D. R. Development and validation of a sociometric instructor evaluation instrument and procedure. Final Report, U. S. Office of Education, Project No. OEC 6-72-0735(509). Washington, D.C.: U. S. Office of Education, 1972. ED 068 077.
20. Hohen, L. P. (Ed.) Teaching behavior improvement program. Detroit, Mich.: Michigan-Ohio Regional Educational Lab., Inc., 1969. ED 034 719.
21. Kapel, D. E. Assessment of a conceptually based instructor evaluation form. Research in Higher Education, 1974, 2, 1-24.
22. Kerlinger, F. N. Student evaluation of faculty professors. School and Society, 1971, 99, 353-356.
23. Klein, S. S. Student influence on teacher behavior. American Educational Research Journal, 1971, 8, 403-422.
24. Lawson, T. E. Formative instructional product evaluation: instruments and strategies. Englewood Cliffs, N.J.: Educational Technology Publications, 1974.
25. Marsh, H. W., Filmer, H., & Thomas, J. S. The validity and usefulness of student evaluations of instructional quality. Journal of Educational Psychology. (In press)
26. Marshall, M. S. Reverse grading. Educational Leadership, 1971, 28, 663-665.
27. McKeachie, W. J. Instructional psychology. In M. R. Rosenzweig and L. W. Porter (Eds.), Annual Review of Psychology. Stanford, Calif.: Annual Reviews, 1974. Pp. 164-193.

28. McKeachie, W. J., Linn, Y., & Mann, W. Student ratings of teacher effectiveness: validity studies. American Educational Research Journal, 1971, 8, 435-445.
29. Miller, M. T. Instructor attitudes toward, and their use of student ratings of teachers. Journal of Educational Psychology, 1971, 62, 235-239.
30. Oles, H. J. & Lencoski, A. Changes in an instructor's self-rating resulting from feedback from student evaluations. Catalog of Selected Documents in Psychology, 1973, 3, 17.
31. Pambookian, H. S. Initial level of student evaluation of instruction as a source of influence on instructor change after feedback. Journal of Educational Psychology, 1974, 66, 52-56.
32. Popham, W. J. Development of a performance test of teaching proficiency. Final Report, U. S. Office of Education, Project No. 5-0566-2-12-1. Washington, D. C.: U. S. Office of Education, 1967. ED 013 242.
33. Rodenwald, H. K., & Carroll, J. L. Problems in use of student ratings of faculty. Perceptual and Motor Skills, 1974, 38, 574.
34. Rous, S. N., et al. The improvement of faculty teaching through evaluation: a follow-up report. Journal of Surgical Research, 1966, 13, 262-266.
35. Sanders, J. R., & Cunningham, D. J. A structure for formative evaluation in product development. Review of Educational Research, 1973, 43, 217-236.
36. Scriven, M. The methodology of evaluation. American Educational Research Association Monograph Series on Curriculum Evaluation, 1967, No. 1, 39-83.
37. Sherman, T. M. The effects of student formative evaluation of instruction on teacher behavior. Unpublished manuscript, Virginia Polytechnic Institute and State University, Blacksburg, Va., 1975.
38. Sherman, T. M., & Winstead, J. C. A formative approach to student evaluation of instruction. Educational Technology, 1975, 15, 34-39.
39. Tolor, A. Evaluation of perceived teacher effectiveness. Journal of Educational Psychology, 1973, 64, 98-104.

40. Trent, J. W., & Conen, A. M. Research on teaching in higher education. In R.M.W. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand McNally and Co., 1973. Pp. 997-1071.
41. Webster, W. J., & Mendro, R. V. The effects of a videotape feedback system on classroom teacher behavior. Journal of School Psychology, 1974, 12, 189-198.
42. Witheiler, P., & Yuher, H. E. Course evaluations at Hofstra University. Hempstead, New York: Hofstra University, Center for the Study of Higher Education, 1970. ED 040 661.
43. Yonge, G. D., & Sassenrath, J. M. Student personality correlates of teacher ratings. Journal of Educational Psychology, 1968, 59, 44-52.
44. Zelby, L. W. Student-faculty evaluation. Science, 1974, 183, 1267-1270.

Items followed by an ED number (for example, ED 050 712) are available from the ERIC Document Reproduction Service (EDRS). Consult the most recent issue of Resources in Education for the address and ordering information.